

Development smart eyeglasses for visually impaired people based on you only look once

Hassan Salam Abdul-Ameer, Hassan Jaleel Hassan, Salma Hameedi Abdullah

Department of Computer Engineering, University of Technology, Baghdad, Iraq

Article Info

Article history:

Received Apr 10, 2021

Revised Dec 22, 2021

Accepted Dec 30, 2021

Keywords:

Computer vision

Convolutional neural network

Object detection

OpenCV

YOLO

ABSTRACT

Visually impaired people are facing many problems in their life. One of these problems is how they can find the objects in their indoor environment. This research was presented to assists visually impaired people in finding the objects in office. Object detection is a method used to detect the objects in images and videos. Many algorithms used for object detection such as convolutional neural network (CNN) and you only look once (YOLO). The proposed method was YOLO which outperforms the other algorithms such as CNN. In CNN the algorithm splits the image into regions. These regions sequentially enters the neural network for object detection and recognition so CNN does not deal with all the regions at the same time but YOLO looks the entire image then it produces the bounding boxes with convolutional network and the probabilities of these boxes, this makes YOLO faster than other algorithms. Open source computer vision (OpenCV) used to capture frames by using camera. Then YOLO used to detect and recognize the objects in each frame. Finally, the sound in Arabic language was generated to tell the visually impaired people about the objects. The proposed system can detect 6 objects and achieve an accuracy of 99%.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Hassan Salam Abdul-Ameer

Department of Computer Engineering, University of Technology

Baghdad, Iraq

Email: ce.19.06@grad.uotechnology.edu.iq

1. INTRODUCTION

The number of blind and visually impaired people is constantly increasing. According to official statistics from the world health organization (WHO), globally, up to the year of 2011, there are about 285 million visually impaired people, 39 million among them are completely blind and 246 million have weak sight [1], [2]. The statistics of WHO in 2018 shows that there is nearly 1 billion blind and visually impaired people [3]. While, in 2020, it became 2.2 billion, this increases the needs for the devices that are used to help the visually impaired people to perform daily tasks. The recent advances in technology lead to develop many devices that are used to assist the visually impaired people such as smart eyeglasses. The proposed smart eyeglasses system was based on computer vision. Computer vision is a technology which has the ability of processing and understanding the photos and videos by using machines [4], [5]. It has many tasks, object detection is one of its fundamental tasks. Object detection is a method that detects the objects in images and videos [6], [7]. It has various applications such as self-driving cars, the applications which are used to help the blind people in recognizing the objects, car plate detection, automated parking systems and face detection [8], [9]. There are many types of deep learning algorithms that are used to perform object detection for instance region based convolutional neural network (R-CNN), and YOLO [10]. The suggested algorithm was you only look once version 3 (YOLO v3). YOLO v3 is based on CNN. CNN is a deep neural network that

consists of one input layer, more than one hidden layer and one output layer. Each layer has different properties. The first layer in Convolution neural network (CNN) is input layer where the image enters the neural network through it. In this layer, the number of neurons is the same as the number of features. The last layer in CNN is the output layer where the number of neurons is the same as the number of classes. Hidden layers are convolution layer, activation layer, pooling layer and fully connected layer. CNN contains at least one convolution layer which computes a dot product between the connected region in the input and the weights to produce the feature map or an activation map. The role of activation layer is to remove the negative values for accelerating the training process. The activation layer result is pooled by pooling layer to simplify the feature map. The fully connected layer is used to connect the outputs from these layers. It is a one-dimensional layer, has all the labels that are to be classified and it produces a score for each label of classification [11]-[13].

The visually impaired people are exposed to many problems in their daily life such as discovering the objects in their environment. The proposed smart eyeglasses system solves this problem by converting the visual scene into voice message. Many researches have been conducted to implement smart eyeglasses and the systems that can be used to help the visually impaired people by using the deep learning algorithms such as R-CNN, CNN, and YOLO. We discuss some of relative methods that can be applied to detect and recognize the objects [14]. Bharti *et al.* [15], implements a system to assist the blind people. CNN, Open-source computer vision (OpenCV), custom dataset and Raspberry Pi are used. The system can detect 16 classes. The accuracy of this system is 90%. Masurekar *et al.* [16], creates an object detection model to help the blind and visually impaired people. YOLO v3 and the custom dataset which contain three classes (bus, mobile and bottle) are used. Sound is generated using Google Text To Speech. They found that the accuracy of this model is 98% and the required time to detect the objects in each image is eight seconds. Vaidya *et al.* [17], Implements an android application and web application for object detection. YOLO v3 with common objects in context (COCO) dataset, are used in this system. They found that the maximum accuracy in mobile phones is 85.5% and 89 % in web applications and the required time is 2 seconds, the time will be increased by increasing the number of objects. Shaikh *et al.* [18], uses Raspberry Pi, YOLO v3 and COCO dataset for implementing an object detection system. The accuracy is 100% (for clock, chair, cellphone and person) and 95% on overall performance. The deep learning algorithms are used to implement other kinds of systems such as a sign language translation and the monitoring systems. Fahad *et al.* [19], implements a sign language translation system. CNN and custom dataset are used. The system converts the sign language into a voice message. Fourty hand gestures are recognized by this system. The achieved accuracy is 98%. Abdulhussein and Raheem [20], implements a hand gesture recognition system using convolutional neural network and custom dataset. Twenty-four letters are recognized. The accuracy is 99.3%. Mahmood and Saud [21], implements a monitoring system for detecting and classifying the moving vehicles in videos using convolutional neural network and the custom dataset. The Accuracy is 92%. Zin *et al.* [22], created a herbal plant recognition system by using convolutional neural network with the custom dataset. Twelve types of plants can be recognized by this system. The accuracy is 99%. Anandhalli *et al.* [23], implements a model for detecting and tracking the vehicle by using convolution neural network and the custom dataset. The achieved accuracy is 90.88%. The proposed smart eyeglasses system uses YOLO v3 with custom dataset. This system produces a high accuracy in detecting and recognizing the objects which is equal to 99%.

2. PROPOSED METHOD

The proposed smart eyeglasses system consists of: Raspberry Pi, USB camera, power bank and earphone. Figure 1 explains the block diagram of the proposed system. The suggested system uses YOLO v3 with the custom dataset for detection and recognition the static objects for indoor environment such as office or room. OpenCV library was used for capturing and processing the images. Play sound library to play the sound from sounds dataset was used with this system. The proposed system was implemented on Raspberry Pi 4 Model B with python language. The proposed method for this smart eyglass system consists of two parts: The first part was the training process of the neural network while the other part was how to use this neural network to detect and recognize the objects. Figure 2 explains the two parts of the proposed method.

The following are the steps for training the proposed deep neural network:

- Step 1: a set of color and high-resolution images with different sizes is collected.
- Step 2: labeling is used to label the objects in each image. Labeling is an image annotation tool which is used for labeling the objects in each image.
- Step 3: an image annotation file was created for each image. The dataset now is ready for training the deep neural network. The training process is executed on graphics processing unit (GPU) of Google Colab for 3000 iterations and takes about three hours.
- Step 4: at the end of the training process the weight file was generated. Figure 2(a), explains the steps for

the training process.

The following are the steps for the second part of proposed method:

- Step 1: the camera will capture the frames (images) by using OpenCV library.
- Step 2: each frame (image) is resized to 416x416 using OpenCV.
- Step 3: YOLO v3 is used to detect and recognize the objects in each image based on weight file.
- Step 4: when the frame has objects, the objects will be detected and recognized. If there is no object in the frame, then the next frame will be selected to detect and recognize the objects in it.
- Step 5: the sound in Arabic language from the sounds dataset will be played by using play sound library to tell the visually impaired people about the objects in the frame. Figure 2(b), explains the object detection and recognition by the proposed system.

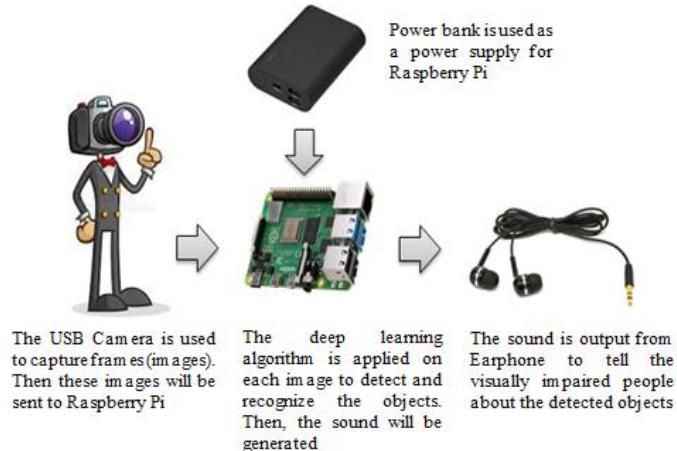


Figure 1. The Block diagram of proposed system

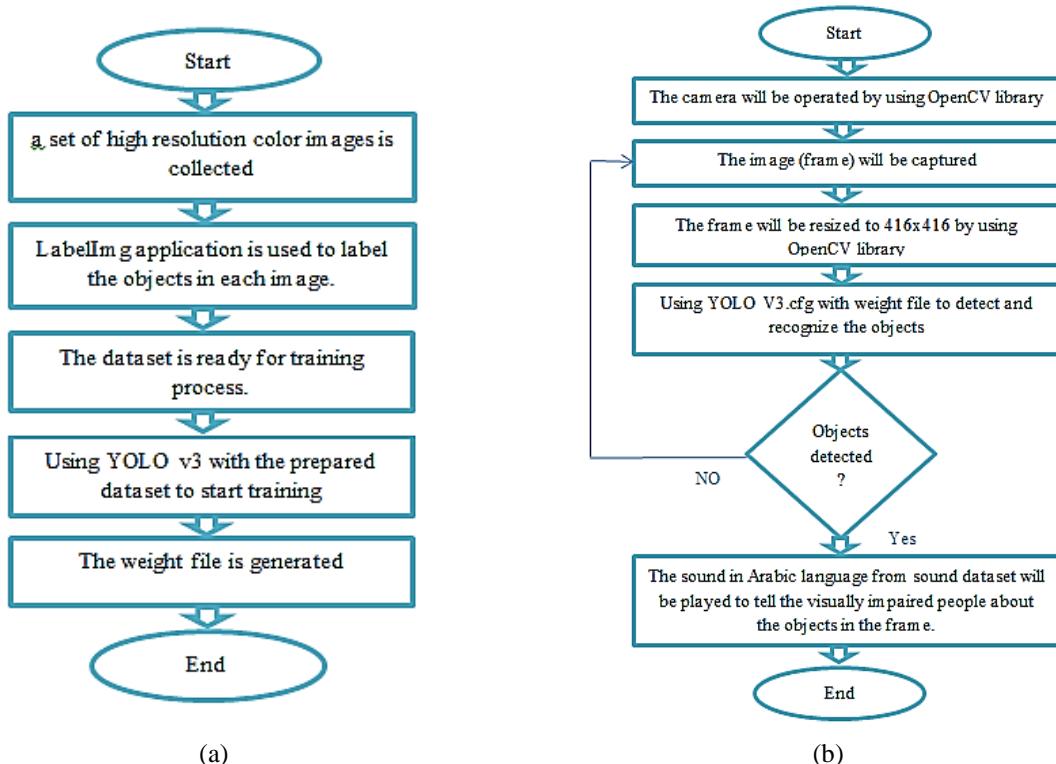


Figure 2. The two parts of proposed method (a) explains the steps for the training process and (b) explains the object detection and recognition by the proposed system

2.1. OpenCV

Open-source computer vision is a library of programming functions which is used for image processing. Image processing is a kind of signal processing in which the input is an image such as a video frame or photograph, the output is an image or set of characteristics related to the image. OpenCV was started as a research project by Intel. It contains various tools to solve computer vision problems. It has low level image processing functions and high level algorithms for detecting faces, feature matching and tracking [24].

2.2. Raspberry Pi 4 model B

Raspberry Pi is the heart of the proposed smart eyeglasses system. Earphone is connected to Raspberry Pi to send sound to the visually impaired people. USB Camera was used because the cable of Raspberry Pi camera is stiff and difficult to maintain. Power bank is the power supply for Raspberry Pi. Raspberry Pi 4 model B with 128 GB SD Card was used in the proposed system. The code in python language was executed on Raspberry Pi.

2.3. YOLO v3

YOLO v3 is consisted of the CNN and an algorithm for processing the output from neural network [25]. CNN is a type of deep neural networks. It is used to process image data. YOLO v3 is a fast, multi-object detection, real-time deep learning algorithm. The excellent processing speed of YOLO v3 is the feature that makes it outperforms the other algorithms, as it can process 45 frames per second. YOLO v3 applies a single CNN to an entire frame, it divides frame into ‘S’ x ‘S’ grids then predicts bounding boxes and finds probabilities for these boxes to detect and recognize the objects [26]. YOLO v3 consists of 106 layers and can detect the objects at three different scales. As shown in Figure 3.

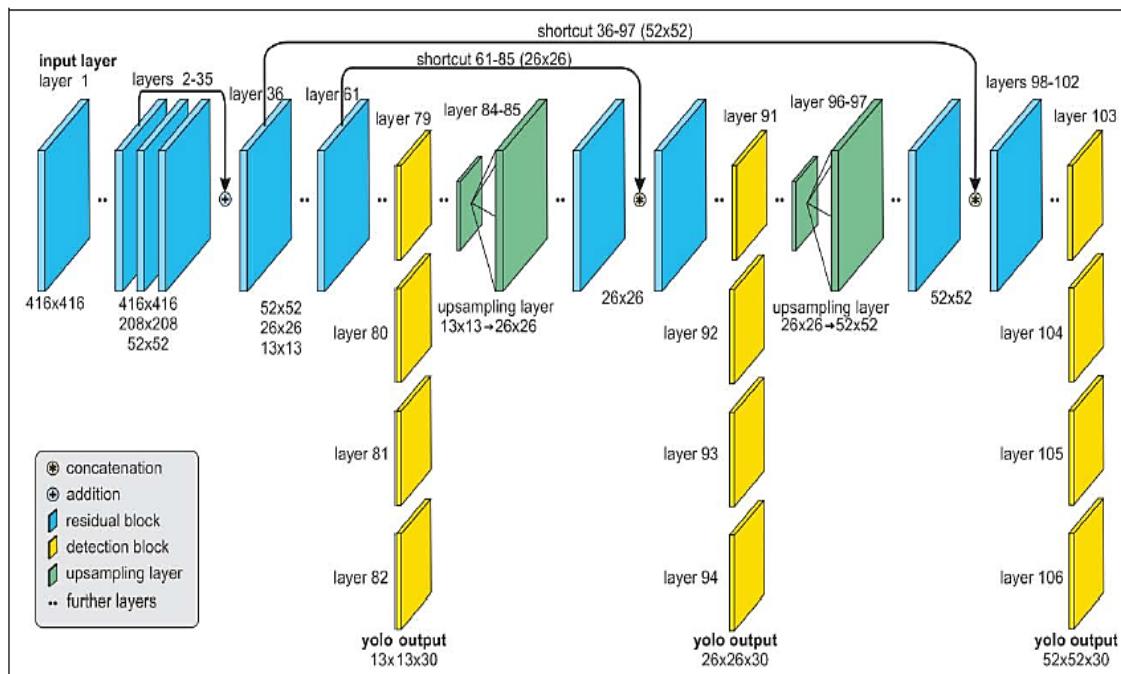


Figure 3. The layers of YOLO v3 [25]

YOLO v3 algorithm works as in the following:

- YOLO v3 takes the images (frames) from camera and analyses each image to detect and recognize the objects in this image. It divides the image into ‘S’ x ‘S’ grids. Each grid has a probability of enclosing multiple or single object. The objects are to be bound by a bounding box. As a result, each grid will have “B” bounding boxes and probabilities of C classes.
- The confidence score is assumed (generally 40% or above) according to which bounding box is predicted against probability of C classes of objects. Predictions with invalid confidence score will not be projected.
- Each boundary box prediction includes 5 values: x, y, w, h, and confidence. The (x, y) is the center of the box, h represents the height and w represents the width. The values of “x”, “y”, “w” and “h” are between [0, 1]. There are 6 class probabilities for each grid cell but only one class probability can be predicted per

cell. The final prediction is of the form $S \times S$ ($B \times 5 + C$). One object only can be detected per cell in the $S \times S$ grid.

- Each grid cell may detect only one object; therefore, anchor box will be used to enable multiple object detection. Consider the image in Figure 4, the midpoints of the car and human are in the same cell; therefore, an anchor box is used. The grid cells in purple color are the anchor boxes for these objects. Any number of the anchor boxes may be used for detecting multiple objects in a single image. Two anchor boxes are used in this image [16].
- When two or more grid cells in $S \times S$ grid have the same object, then the object's center point is determined and the cell which has this center point is selected. There are two methods for dealing with multiple bounding boxes that are found around the objects. The methods are non-max suppression (NMS) and intersection over union (IoU). In IoU method, If the intersection over union value for the bounding boxes is equal to or greater than the threshold value (our threshold value is 0.5) then prediction is good. The accuracy will increase by increasing the threshold value [16]. In the second method (NMS), the boxes which have high probability will be taken and the boxes with high IoU will be suppressed. This process is repeated until a box is selected and considered as the bounding box for the object [10].
- As previously mentioned, YOLO v3 consists of 106 layers. Our proposed method consists of 94 layers instead of 106 layers, by removing the last 12 layers from YOLO v3 algorithm to decrease the required time for object detection while maintain the accuracy as our system does not deal with very small objects, but with large and medium objects to enable the blind people to discover the objects in front of them.



Figure 4. The anchor boxes

2.4. Custom dataset (image dataset)

In the training process for neural networks, many images are required to train the deep learning model. The prepared dataset for the suggested system consists of 1560 labeled images for 6 objects (TV, bottle, person, chair, laptop and table). The number of images that belongs to TV was 180 images, for person was 600, 180 for bottle, 180 for chair, 220 for table and 200 for laptop. The images are in different sizes. These images are in .JPG format. Figure 5 explains some of dataset images.



Figure 5. Image dataset

2.5. Sound dataset

A set of voice messages in Arabic language is created and stored in the Raspberry Pi. When the object is detected and recognized then the sound will be played by using the play sound library, to tell the visually impaired people about the objects in each frame (in front of him/her). The voice messages are in .MP3 format. This method will convert the text into voice message without using the Internet and at high speed.

3. RESULTS AND ANALYSIS

YOLO v3 which is an accurate and fast algorithm was used in the proposed system. The dataset which consisted of 1560 colour image was splitted into two groups. The first contains eighty-five percent of the total images as training images while the second contains the remainder as testing images. The training process is executed on GPU of Google Colab and takes about three hours. The deep neural network is trained for (3000) iterations. The proposed smart eyeglasses system can be used for indoor environments such as room or office and it can detect multiple objects. Also, the system can detect the objects even if the distance between the USB Camera and objects is greater than 3 meters. The mean average precision (mAP) of the smart eyeglasses system was 100%. The mAP is used for evaluating the performance of the suggested system. Figure 6 shows the mAP and Loss for the suggested method. There are other values that are used for evaluating the performance of the suggested system such as precision, IoU, Recall, F1-score, true positive (TP), false positive (FP) and false negative (FN). Figure 7 explains the results of training process (the performance of the proposed method). Figure 8 shows the object detection by using smart eyeglasses. Table 1 explains the comparison between suggested approach and other related approach.

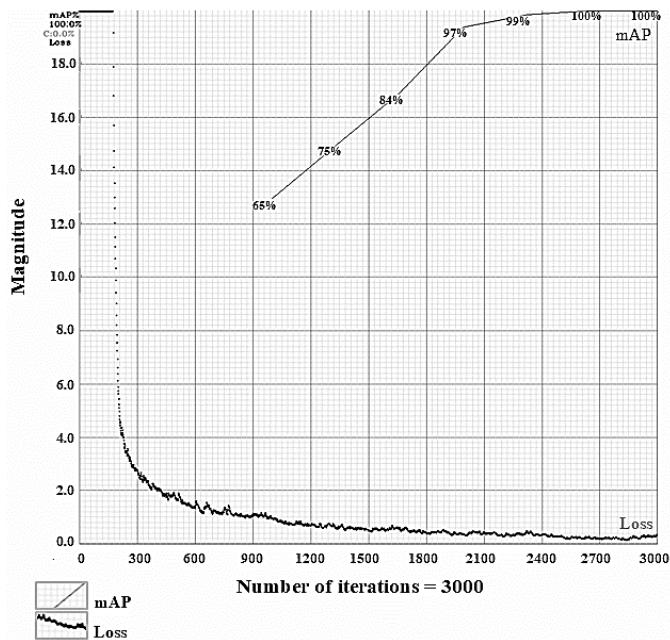


Figure 6. Loss and mAP

```

class_id = 0, name = TV, ap = 100.00%      (TP = 27, FP = 0)
class_id = 1, name = Person, ap = 100.00%     (TP = 90, FP = 0)
class_id = 2, name = bottle, ap = 100.00%     (TP = 27, FP = 0)
class_id = 3, name = chair, ap = 100.00%      (TP = 27, FP = 1)
class_id = 4, name = table, ap = 100.00%      (TP = 33, FP = 0)
class_id = 5, name = laptop, ap = 100.00%     (TP = 30, FP = 1)

for conf_thresh = 0.25, precision = 0.99, recall = 1.00, F1-score = 1.00
for conf_thresh = 0.25, TP = 234, FP = 2, FN = 0, average IoU = 86.48 %

IoU threshold = 50 %, used Area-Under-Curve for each unique Recall
mean average precision (mAP@0.50) = 1.000000, or 100.00 %

```

Figure 7. The result of training process

3.1. Confusion matrix

The confusion matrix, which is also called an error matrix, is a summary that give the result of the prediction. The number of incorrect and correct predictions is summarized with counted values and broken down class by class. The error matrix (confusion matrix) explains how the model is confused when it makes predictions. The confusion matrix gives the insight not only into the errors being made by the classifier but

also give the type of error [16]. FP refers to the number of incorrect detections. The number of correctly detected objects is represented by TP. The FN refers to number of missed detection.

For TV	TP = 27 and FP = 0	For chair	TP = 27 and FP = 1	Total TP = 234
For person	TP = 90 and FP = 0	For table	TP = 30 and FP = 0	Total FP = 2
For bottle	TP = 27 and FP = 0	For laptop	TP = 33 and FP = 1	

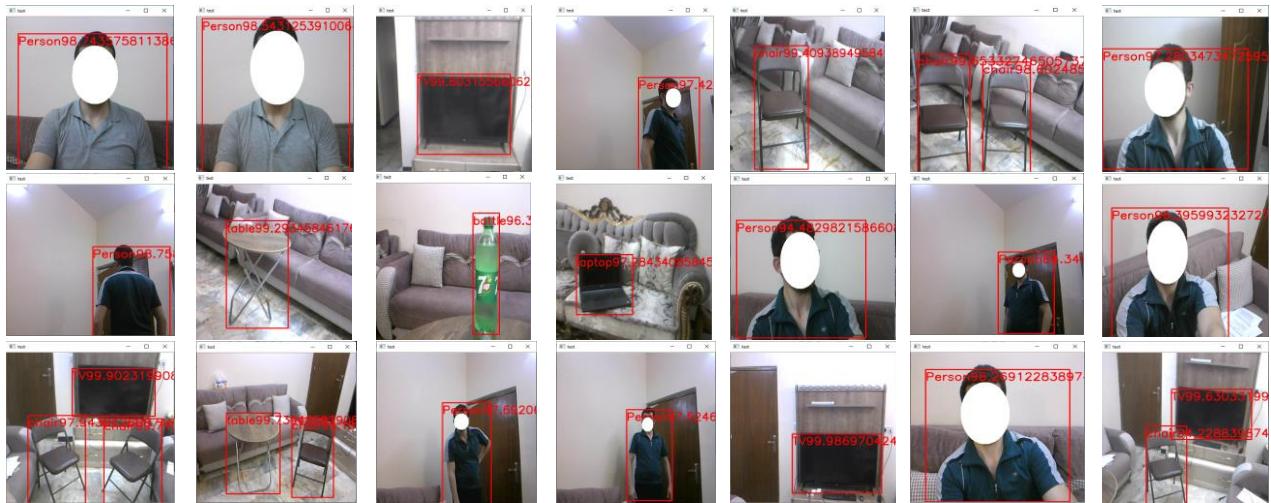


Figure 8. The results of object detection using smart eyeglasses system

Table 1. The comparison between the suggested approach and other related approach

Author	Method	Accuracy	no. of objects	Processing time
Masurekar <i>et al.</i> [17]	YOLO v3, Custom dataset and Google Text to Speech (GTTS)	98%	3	8 sec
The proposed method	YOLO v3, Custom Dataset, OpenCV, Play sound and Raspberry Pi.	99%	6	Nearly 20 sec

3.2. Precision

The precision is used to measure how accurate the predictions are. Precision calculation will be as: the division of TP over the sum of FP and TP. In (1) explains the precision [27]. The obtained value is 0.99.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (1)$$

3.3. Recall

The Recall is used for calculating the true prediction from the all correctly predicted data. Recall calculation will be as: the division of TP over the sum of FN and TP. In (2) explains the recall metric [27]. The obtained Recall value is 1.00.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

3.4. F1-score

The harmonic mean (HM) of the Precision and the Recall. It is one of metrics that used for performance evaluation. Obtained value is 1.00.

3.5. IoU

The division of intersection area over area of union between ground truth bounding box and detection bounding box for a specific threshold. In (3) shows the IoU [27]. The obtained average IoU is 86.48%.

$$\text{IoU} = \frac{\text{Area of Intersection}}{\text{Area of Union}} \times 100\% \quad (3)$$

3.6. mAP

The mean average precision is the average of average precision (AP) that is calculating for all classes. In (4) explains the mean average precision. The obtained mAP is 100%.

$$mAP = \frac{\text{sum of AP for the total classes}}{\text{no.of total classes}} \times 100\% \quad (4)$$

4. CONCLUSION

There are many types of algorithms that are used for object detection and recognition for instance R-CNN, fast R-CNN and YOLO. The suggested smart eyeglasses system uses YOLO v3 because it is a fast and accurate method. YOLO v3 can detect multiple objects in an image. The play sound is a library that is used to play the sound so the suggested system not needs connection to the internet for converting the text into a voice message. The proposed method achieves an accuracy of 99%. The required time for detection is nearly twenty seconds on Raspberry Pi while on the personal computer is nearly two seconds. In future, the smart eyeglasses can be implementing on NVIDIA Jetson Nano instead of Raspberry Pi to decrease the required time to detect the objects for each frame. Gender prediction technique and age prediction technique can also be added to predict the gender and age if the detected object is person. In addition, the technique of face recognition can be added to assist the blind people in recognizing the people in front of them.

REFERENCES

- [1] J. Bai, S. Lian, Z. Liu, K. Wang, and D. Liu, "Smart guiding glasses for visually impaired people in indoor environment," *IEEE Transactions on Consumer Electronics*, vol. 63, no. 3, pp. 258–266, 2017, doi: 10.1109/TCE.2017.014980.
- [2] H. Jabnoun, F. Benzarti, and H. Amiri, "Object detection and identification for blind people in video scene," *International Conference on Intelligent Systems Design and Applications (ISDA)*, vol. 2016-June, pp. 363–367, 2016, doi: 10.1109/ISDA.2015.7489256.
- [3] N. Satani, S. Patel, and S. Patel, "AI Powered Glasses for Visually Impaired Person," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 9, no. 2, pp. 416–421, 2020, doi: 10.35940/ijrte.b3565.079220.
- [4] H. Bhorshetti, S. Ghuge, A. Kulkarni, P. S. Bhingarkar, and P. N. Lokhande, "Low Budget Smart Glasses for Visually Impaired People," *10th International Conference on Intelligent Systems and Communication Networks (IC-ISCN 2019)*, 2019, pp. 48–52.
- [5] M. Vaidhehi, V. Seth, and B. Singhal, "Real-Time Object Detection for Aiding Visually Impaired using Deep Learning," *International Journal of Engineering and Advance Technology (IJEAT)*, vol. 9, no. 4, pp. 1600–1605, 2020, doi:10.35940/ijeat.d8374.049420.
- [6] V. Kharchenko and I. Chyrka, "Detection of Airplanes on the Ground Using YOLO Neural Network," *International Conference Mathematical Methods Electromagnetic Theory, MMET*, 2018, pp. 294–297, doi: 10.1109/MMET.2018.8460392.
- [7] Z. Cheng, J. Lv, A. Wu, and N. Qu, "YOLOv3 Object Detection Algorithm with Feature Pyramid Attention for Remote Sensing Images," *Sensors Mater.*, vol. 32, no. 12, pp. 4537–4558, 2020, doi: 10.18494/sam.2020.3130.
- [8] S. N. Srivatsa, G. Sreevaths, G. Vinay, and P. Elaiyaraaja, "Object Detection using Deep Learning with OpenCV and Python," *International Research Journal of Engineering and Technology (IRJET)*, vol. 8, no. 1, pp. 227–230, 2021, [Online]. Available: <https://www.irjet.net/archives/V8/i1/IRJET-V8I145.pdf>
- [9] M. Shah and R. Kapdi, "Object detection using deep neural networks," in *Proc. 2017 International Conference on Intelligent Computing and Control Systems (ICICCS) 2017*, 2017, pp. 787–790, doi: 10.1109/ICCONS.2017.8250570.
- [10] S. Geethapriya, N. Duraimurugan, and S. P. Chokkalingam, "Real time object detection with yolo," *International Journal of Engineering and Advanced Technology (IJEAT)*, vol. 8, no. 3, pp. 578–581, 2019.
- [11] K. Poidar, C. D. Pai, and S. Akolkar, "A Convolutional Neural Network based Live Object Recognition System as Blind Aid," *arXiv*, 2018. [Online]. Available: <https://arxiv.org/pdf/1811.10399.pdf>
- [12] P. S. Bhairnallykar, A. Prajapati, A. Rajbhar, and S. Mujawar, "Convolutional Neural Network (CNN) for Image Detection," *International Research Journal of Engineering and Technology (IRJET)*, vol. 7, no. 11, pp. 1239–1243, 2020. [Online]. Available: <https://www.irjet.net/archives/V7/i11/IRJET-V7I11204.pdf>.
- [13] T. S. Gunawan *et al.*, "Development of video-based emotion recognition using deep learning with Google Colab," *TELKOMNIKA Telecommunication, Computing, Electronics and Control*, vol. 18, no. 5, pp. 2463–2471, 2020, doi:10.12928/TELKOMNIKA.v18i5.16717.
- [14] A. Abdurasyid, I. Indrianto, and R. Arianto, "Detection of immovable objects on visually impaired people walking aids," *TELKOMNIKA Telecommunication, Computing, Electronics and Control*, vol. 17, no. 2, pp. 580–585, 2019, doi: 10.12928/TELKOMNIKA.V17I2.9933.
- [15] R. Bharti, K. Bhadane, P. Bhadane, and A. Gadhe, "Object Detection and Recognition for Blind Assistance," *International Research Journal of Engineering and Technology (IRJET)*, vol. 6, no. 5, pp. 7085–7087, 2019. [Online]. Available: <https://www.irjet.net/archives/V6/i5/IRJET-V6I51002.pdf>.
- [16] O. Masurekar, O. Jadhav, P. Kulkarni, and S. Patil, "Real Time Object Detection Using YOLOv3," *Int. Res. J. Eng. Technol.*, vol. 7, no. 3, pp. 3764–3768, 2020, [Online]. Available: <https://www.irjet.net/archives/V7/i3/IRJET-V7I3756.pdf>
- [17] S. Vaidya, N. Shah, N. Shah, and R. Shankarmani, "Real-Time Object Detection for Visually Challenged People," in *Proc. International Conference on Intelligent Computing and Control Systems (ICICCS 2020)*, 2020, pp. 311–316, doi: 10.1109/ICICCS48265.2020.9121085.
- [18] S. Shaikh, V. Karale, and G. Tawde, "Assistive Object Recognition System for Visually Impaired," *International Journal of Engineering Research & Technology (IJERT)*, vol. 9, no. 9, pp. 736–740, 2020, doi: 10.17577/ijertv9is090382.
- [19] A. A. Fahad, H. J. Hassan, and S. H. Abdullah, "Deep Learning-based Deaf & Mute Gesture Translation System," *International Journal of Science and Research (IJSR)*, vol. 9, no.5 , pp. 288–292, 2020, doi: 10.21275/SR20503031800.

- [20] A. A. Abdulhussein and F. A. Raheem, "Hand Gesture Recognition of Static Letters American Sign Language (ASL) Using Deep Learning," *Engineering and Technology Journal*, vol. 38, no. 6, pp. 926–937, 2020, doi: 10.30684/etj.v38i6a.533.
- [21] S. S. Mahmood and L. J. Saud, "An Efficient Approach for Detecting and Classifying Moving Vehicles in a Video Based Monitoring System," *Engineering and Technology Journal*, vol. 38, no. 6A, pp. 832–845, 2020, doi: 10.30684/etj.v38i6a.438.
- [22] I. A. M. Zin, Z. Ibrahim, D. Isa, S. Aliman, N. Sabri, and N. N. A. Mangshor, "Herbal plant recognition using deep convolutional neural network," *Bulletin of Electrical Engineering and Informatics (BEEI)*, vol. 9, no. 5, pp. 2198–2205, 2020, doi: 10.11591/eei.v9i5.225.
- [23] M. Anandhalli, V. P. Baligar, P. Baligar, P. Deepsr, and M. Iti, "Vehicle detection and tracking for traffic management," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 10, no. 1, pp. 66–73, 2021, doi: 10.11591/ijai.v10.i1.pp66-73.
- [24] M. Naveenkumar and V. Ayyasamy, "OpenCV for Computer Vision Applications," in *Proc of National Conference on Big Data and Cloud Computing (NCBDC'15)*, 2015, pp. 52–56, [Online]. Available: https://www.researchgate.net/publication/301590571_OpenCV_for_Computer_Vision_Applications.
- [25] A. Corovic, V. Ilic, S. Duric, M. Marijan, and B. Pavkovic, "The Real-Time Detection of Traffic Participants Using YOLO Algorithm," in *Proc. 2018 26th Telecommunications Forum, TELFOR 2018*, 2018, doi: 10.1109/TELFOR.2018.8611986.
- [26] M. Burić, M. Pobar, and M. Ivašić-Kos, "Adapting yolo network for ball and player detection," in *Proc. of the 8th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2019)*, 2019, pp. 845–851, doi: 10.5220/0007582008450851.
- [27] Y. Huang, J. Zheng, S. Sun, C. Yang, and J. Liu, "Applied sciences Optimized YOLOv3 Algorithm and Its Application in Traffic Flow Detections," *Applied Sciences*, 2020.

BIOGRAPHIES OF AUTHORS



Hassan Salam Abdul-Ameer received his Bachelor of Computer Engineering from University of Technology (UOT), Iraq, in 2015. He is currently working on his Master degree at University of Technology. His area of interest are computer vision, deep learning. He can be contacted at email: ce.19.06@grad.uotechnology.edu.iq



Asst. professor Hassan J. Hassan is currently a head of department of computer engineering, university of technology (UOT), Iraq, since 2019. He obtained his PhD in computer network from control and systems engineering department, university of technology in 2013. His research interest includes, wireless communication networks, Internet of things (IoT), Network congestion control, Advanced image processing. He can be contacted at email: 60012@uotechnology.edu.iq



Dr. Salma Hameedi Abdullah is currently a lecturer at the faculty of Computer Engineering Department, University of Technology, Iraq. She obtained her Ph.D. in computer vision from the University of Technology in 2018. Her research interest includes computer vision, deep learning, internet of things (IoT), and pattern recognition. She can be contacted at email: 120015@uotechnology.edu.iq